



Block Pickard Models for Two-Dimensional Constraints

Forchhammer, Søren; Justesen, Jørn

Published in:

IEEE Transactions on Information Theory

Link to article, DOI:

[10.1109/TIT.2009.2027505](https://doi.org/10.1109/TIT.2009.2027505)

Publication date:

2009

Document Version

Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Forchhammer, S., & Justesen, J. (2009). Block Pickard Models for Two-Dimensional Constraints. *IEEE Transactions on Information Theory*, 55(10), 4626-4634. <https://doi.org/10.1109/TIT.2009.2027505>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Block Pickard Models for Two-Dimensional Constraints

Søren Forchhammer, *Member, IEEE*, and Jørn Justesen, *Member, IEEE*

Abstract—In Pickard random fields (PRF), the probabilities of finite configurations and the entropy of the field can be calculated explicitly, but only very simple structures can be incorporated into such a field. Given two Markov chains describing a boundary, an algorithm is presented which determines whether a PRF consistent with the distribution on the boundary and a 2-D constraint exists. Iterative scaling is used as part of the algorithm, which also determines the conditional probabilities yielding the maximum entropy for the given boundary description if a solution exists. A PRF is defined for the domino tiling constraint represented by a quaternary alphabet. PRF models are also presented for higher order constraints, including the no isolated bits (n.i.b.) constraint, and a minimum distance 3 constraint by defining super symbols on blocks of binary symbols.

Index Terms—Pickard random fields, minimum distance 3 constraint, n.i.b. constraint, two-dimensional (2-D) capacity, 2-D constraints, 2-D entropy.

I. INTRODUCTION

WE consider two-dimensional (2-D) fields specified by shift invariant constraints of finite extent (N, M) over some finite alphabet \mathcal{A} . A constraint is defined by a list, \mathcal{F} , of forbidden configurations of symbols in \mathcal{A} . Each forbidden configuration is contained within a rectangle of maximum size $N \times M$. A configuration on a finite segment of the plane containing no forbidden configurations is called an admissible configuration.

Let $E(n, m)$ be the set of admissible configurations on an $n \times m$ rectangle for a given field F . The combinatorial entropy (or capacity) of F is defined as

$$C(F) = \limsup_{n, m \rightarrow \infty} \frac{\log_2 |E(n, m)|}{mn}. \quad (1)$$

As discussed in [1], this limit is well defined, even though it may not be computable. The limit is identical if we consider rotated rectangles, parallelograms, or other segments where the length of the boundary is small compared to the area [2]–[4].

In this paper, we shall approach the entropy by assigning probability measures to the configurations. Let μ_E be a probability measure on $\mathcal{A}^{n \times m}$ that agrees with the constraint on an $n \times m$ rectangle. That is, each of the $|\mathcal{A}|^{n \times m}$ possible configurations, \mathbf{x} , that contains forbidden words in \mathcal{F} , have probability

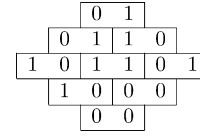


Fig. 1. A n.i.b. configuration on 3×3 blocks.

zero under μ_E , $0 \leq \mu_E \leq 1$. Let the elements of the $n \times m$ configuration be denoted x_{ij} , $1 \leq i \leq n$, $1 \leq j \leq m$. We shall also use $\mu_{n \times m}$ for the measure on a $n \times m$ rectangle. The (measure theoretic) entropy of μ_E is defined as

$$H(\mu_E) = -\frac{1}{nm} \sum_{\mathbf{x} \in \mathcal{A}^{n \times m}} \mu_E(\mathbf{x}) \log_2 \mu_E(\mathbf{x}). \quad (2)$$

A measure is said to be (locally) stationary if two configurations within the rectangle that differ only by a translation, i.e., adding a constant pair (i', j') to the indices, have the same probability. For finite (n, m) it is a standard result that $H(\mu_E) \leq \log(|E(n, m)|)$, and that equality holds if all admissible outcomes have the same probability. Thus if the entropy converges for (n, m) going to infinity, it is a lower bound on the capacity. Besides the entropy, a stationary measure also provides useful information as correlation properties and the spectrum.

So far analytical results are largely limited to first-order (2×2) constraints, and even such cases are in general extremely difficult. We analyze some cases of larger constraints by considering rectangular blocks of given size covering the plane. The approach is illustrated by the following example.

Example 1.1: Consider the “no isolated bits” (n.i.b.) 2-D constraint [5] characterized by the two forbidden configurations:

$$\mathcal{F} = \left\{ \begin{array}{ccc} 0 & & 1 \\ 0 & 1 & 0, 1 \\ 0 & & 1 \end{array} \right\}.$$

Introducing blocks of 1×2 elements, the plane may be covered by blocks, where in each row the blocks are offset by one element (Fig. 1).

Section II describes finite causal models in general, and Pickard random fields (PRF) [6] are reviewed. They are defined by probability distributions conditioned on a finite part of the “past”. In Section III, we consider PRF over large alphabets and approach the problem starting with two Markov chains specifying rows and columns. It is shown that the transition matrices of the Markov chains must commute. If this is satisfied, a method based on iterative scaling for optimizing the entropy of the PRF for a constrained field is presented. In Section IV, we define block based Pickard random fields and consider the application to constrained fields. Cases discussed include the

Manuscript received July 13, 2007; revised May 19, 2009. Current version published September 23, 2009.

The authors are with the DTU Fotonik, B.343, Technical University of Denmark, DK-2800 Lyngby, Denmark (e-mail: sofo@fotonik.dtu.dk; joju@fotonik.dtu.dk).

Communicated by T. Etzion, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2009.2027505

n.i.b. constraint and a field with minimum distance 3 between 1's.

II. CAUSAL MODELS AND PICKARD FIELDS

In a MRF, the conditional probability of a finite segment of the plane is specified conditioned on the boundary, which may be chosen wide enough to ensure that the constraint is satisfied. In particular, we can assign equal probabilities to all admissible outcomes within the segment, and it is clear that this assignment is consistent with assignments on subsets. If the distribution converges to a limit when the size of the segment increases (thus becoming independent of the boundary), it follows that the entropy reaches the combinatorial entropy. However there is no practical way of finding the probabilities for large segments.

A. Causal Models

The boundary may be chosen to extend across the whole plane and the conditional probability of a configuration given the outcome in the upper half plane can be specified to depend only on the bottom $N-1$ rows. When a two-dimensional field is specified by a shift invariant constraint of finite extent, the value of a symbol at position (i', j') is constrained by values in a finite neighborhood, $i' - N < i < i' + N, j' - M < j < j' + M$. For practical reasons the field is often processed by taking the symbols in some particular order, usually one row at a time from the top, each row being read from left to right. In this way, we can talk about the symbols preceding (i', j') , i.e., $(i < i')$ or $(i = i'$ and $j < j')$, as the past set.

Definition 2.1: A causal model describes the symbol in position (i', j') in terms of a finite set of symbols belonging to the past set.

Thus if we neglect a border of $N-1$ rows and $M-1$ columns, a causal model gives a shift invariant description of the field. However, in general the conditional probability of a symbol in a MRF depends on an infinite part of the past.

We are interested in describing probability distributions that allow us to express the conditional distribution of a symbol as depending on a finite number of variables, since this will allow us to calculate the relevant entropy, correlation functions and spectrum as well as to study methods for encoding information into the field or compressing an outcome of the field interpreted as a source. The only types of nontrivial 2-D fields which satisfy this requirement so far are variations of Pickard fields [6], [7].

B. Pickard Random Fields

A Pickard random field is described by the probability distribution of four symbols [6]: Let A, B, C, D be random variables over \mathcal{A} in a 2×2 rectangle with relative positions

$$\begin{array}{cc} A & B \\ C & D \end{array}$$

Let X, Y, Z be random variables and let $X \perp Y \mid Z$ denote that X and Y are independent given Z . We assume the independence condition $B \perp C \mid A$ for the models under consideration.

Consider $\mu_{2 \times 2}$ defined by the joint distribution $(ABCD)$ of the variables. Likewise, let for a subset of variables, e.g., (ABC)

denote the joint distribution of the variables and $(Y|X)$ the conditional probability distribution of Y given X . The probabilities of $(ABCD)$ are expressed by

$$P(ABCD) = P(D|ABC)P(ABC) \quad (3)$$

and due to the PRF independence condition $B \perp C \mid A$

$$P(ABC) = P(B|A)P(C|A)P(A). \quad (4)$$

Each symbol has the probability distribution (A) and further the model is specified by the three conditional probability distributions $(B|A), (C|A)$ and $(D|ABC)$. Derived from $\mu_{2 \times 2}$, a measure, $\mu_{n \times m}$ [6], on an $n \times m$ rectangle is defined by the distribution on the boundary \mathbf{x}_δ

$$P(\mathbf{x}_\delta) = P(A = x_{11}) \prod_{i=2}^n P(C = x_{i1} | A = x_{i-11}) \times \prod_{j=2}^m P(B = x_{1j} | A = x_{1j-1}) \quad (5)$$

and the distribution of \mathbf{x} conditioned on the boundary \mathbf{x}_δ

$$P(\mathbf{x} | \mathbf{x}_\delta) = \prod_{i=2}^n \prod_{j=2}^m P(D = x_{ij} | A = x_{i-1j-1}, B = x_{i-1j}, C = x_{ij-1}). \quad (6)$$

Thus, the conditional distributions $(B|A)$ and $(C|A)$ define two Markov chains on the boundary (5). Let \mathbf{R} denote the transition matrix for the Markov chain left-to-right given by $(B|A)$ and \mathbf{S} denote the transition matrix for the Markov chain from the bottom and up, i.e., given by $(A|C)$, which is the reversed Markov chain of $(C|A)$.

To ensure stationarity when extending to the $n \times m$ rectangle, we follow [6] and note that $B \perp C \mid D$ provides a solution. This is expressed by the following theorem based on Pickard ([6, Corollary of Theorem 2]).

Theorem 2.2: Let $\mu_{2 \times 2}$ be a stationary measure induced by $(ABCD)$, satisfying $B \perp C \mid A$. Then for all $(n, m), \mu_{n \times m}$ (5)–(6) is stationary if $\mu_{2 \times 2}$ satisfies either

$$B \perp C \mid D \quad (7)$$

or

$$(A \perp D \mid B \text{ and } A \perp D \mid C). \quad (8)$$

We shall refer to the fields defined by Theorem 2.2 as Pickard random fields (PRF). The pair of independence conditions, (4) and (7), lead to an expression of the probability distribution of the form (5)–(6). The second pair of conditions (8) simply comes from the first by symmetry [7]. [Actually (8), without requiring $B \perp C \mid A$, are sufficient for a stationary measure on an $n \times m$ rectangle. The description of this unilateral measure starts in the upper right hand corner, x_{1m} , instead of the upper left hand corner, x_{11} , as in (5)]. By the stationarity of the Markov chains, any row (or column) has the same distribution as the first row (or column) and further any $s \times t$ rectangle has the same distribution, namely $\mu_{s \times t}$. Pickard [6] also showed

that any s consecutive rows (or columns) form a Markov chain. For any pair of variables, U and V , in the plane, we can calculate the joint distribution (UV) by repeated application of the independence conditions. From these distributions we may derive the correlation functions and power spectrum of fields with real alphabets.

Let $H(D|ABC)$ denote the conditional entropy of D given ABC .

Theorem 2.3: The entropy per symbol (2) of a stationary measure $\mu_E = \mu_{n \times m}$ defined by Theorem 2.2 is bounded by

$$H(\mu_E) \geq H(D|ABC). \quad (9)$$

Proof: The $n \times m$ elements, \mathbf{x} , of the measure $\mu_{n \times m}$ may be divided into the boundary, \mathbf{x}_δ , (5) and the remaining $\mathbf{x} \setminus \mathbf{x}_\delta$ (interior) elements (6). The entropy of each element of the latter is given by $H(D|ABC)$ due to the stationarity and the chain rule. The entropy of the boundary is not less than this as the distribution on the boundary is given by (A) and the Markov chains ($B|A$) and ($C|A$), which are identical to the marginal distribution on the rows (and columns) of the remaining interior, due to the stationarity. \square

Thus $H(D|ABC)$ is a lower bound for $H(\mu_E)$ and the capacity of the constraint, $C(F)$. Clearly, a Pickard field is a special case of an MRF [6], [7].

III. CONSTRUCTING PICKARD RANDOM FIELDS

For larger constraints and causal models with causal neighborhood extending beyond the three closest causal variables, we shall use blocks of symbols as in the construction of Example 1.1. Thus several symbols are collected into a block having k elements, which is treated as a super symbol drawn from the extended finite alphabet, \mathcal{A}^k , of size m and the plane is covered with such blocks arranged in a suitable grid. First this is addressed by considering construction of constrained Pickard random fields over an m -ary alphabet. Thereafter in Section IV, this is used for construction of block Pickard fields.

A. Constructing Pickard Random Fields From Two Markov Chains

We consider the configurations on symbols over a finite m -ary alphabet and proceed by considering the construction of Pickard random fields over a finite alphabet, taking the two Markov chains of the boundary (5) as point of departure instead of the stationary distribution ($ABCD$) as in Section II.

It is essential for the construction that the stationary distribution on the states is the same for the two Markov chains. However, as discussed above, the pairs AB and AC may have different distributions. Nevertheless, we can restate the Pickard independence assumptions (4) and (7) as

$$P(B|AC) = P(B|A), P(C|BD) = P(C|D) \quad (10)$$

$$P(C|AB) = P(C|A), P(B|CD) = P(B|D) \quad (11)$$

and again conclude that they imply that the boundaries are Markov chains.

Given the two Markov chains of the boundary (5), the distribution of BC can then be found in two ways by considering the triples CAB and CDB .

We consider the two transition matrices for the boundary Markov chains (5), \mathbf{R} and \mathbf{S} . Given the distribution of C , defined by a vector, we can find the distribution of B by multiplying by either \mathbf{RS} or \mathbf{SR} . Since the conditional distribution of B is defined conditioned on any value of C , the products have to be identical column by column, and we conclude that the matrices must commute

$$\mathbf{RS} = \mathbf{SR}. \quad (12)$$

This constraint can be written out as a set of nonlinear equations in the entries of \mathbf{R} and \mathbf{S} . However, it is often useful to consider the eigenvectors of \mathbf{R} and \mathbf{S} . If $\mathbf{R}v = \lambda v$, it follows that

$$\mathbf{SR}v = \lambda \mathbf{S}v = \mathbf{RS}v$$

and consequently $\mathbf{S}v$ is also an eigenvector of \mathbf{R} with the same eigenvalue. In our constructions we shall not consider the special cases of multiple eigenvalues, but require that the matrices have the same eigenvectors. In particular, this means that the two processes have the same stationary distributions.

The existence of solutions can often be decided by considering periodic patterns satisfying the given constraints. If such a pattern can be described by a pair of deterministic transition matrices, it follows that the equations have at least this type of solutions.

We shall only consider irreducible Markov chains. If a chain were not irreducible, we should reduce it to an irreducible component. Once a pair of (irreducible) MCs with commuting transition matrices, \mathbf{R}, \mathbf{S} , have been obtained, the final step is to get the distribution ($ABCD$) and thereby $P(D|ABC)$ consistent with the distributions on the triples, (CAB) and (CDB), as defined by \mathbf{R} and \mathbf{S} . By (12), the marginal distribution on (BC) is identical for (CAB) and (CDB) and these triples define $P(A|BC)$ and $P(D|BC)$. The requirement for $P(AD|BC)$ may be formulated by a system of linear equations for each pair of values for $BC = bc$

$$P(A = i|bc) = \sum_{d \in \mathcal{A}} P(A = i, D = d|bc), \quad i \in \mathcal{A} \quad (13)$$

$$P(D = j|bc) = \sum_{a \in \mathcal{A}} P(A = a, D = j|bc), \quad j \in \mathcal{A}. \quad (14)$$

Again there are clearly more variables than equations, but the equations may not have a (positive) solution. At this stage we shall limit the terms of the distribution ($ABCD$) to those configurations that satisfy the constraints, i.e., the PRF shall assign probability $P(abcd) = 0$ to forbidden configurations, $abcd$.

1) *Tiling With Dominoes Using PRF:* We describe domino tiling of the plane as a Pickard field. The combinatorial entropy of this field is known explicitly [8], [9], and the numerical value is approximately 0.42. Nevertheless certain configurations of dominoes give rise to long range effects, and it is not easy to get a random field with an entropy that approaches the theoretical value. Below a very simple field with a modest entropy is constructed.

The dominoes are represented as occupying two adjacent positions in a rectangular grid, and depending on the orientation, they are labelled L (eft), R (ight), U (p), and W (doWn). Clearly an L is always followed by an R in a row, and an U by a W in a column.

The boundary is described by two Markov chains, one for rows (left to right having transition matrix \mathbf{R}) and one for the diagonals from the lower left to upper right (having transition matrix \mathbf{S}). (The same orientation of the Markov chains was applied to the binary hard-square model in [4].) A state in each chain is a single symbol, and for the construction of the field we consider the configuration.

$$\begin{pmatrix} - & A & B \\ C & D & - \end{pmatrix}.$$

Assuming that the states have the same stationary probability, we can write the row Markov chain transition matrix with transitions from columns to rows as

$$\mathbf{R} = \begin{pmatrix} 0 & b+2c+d-1 & 1-b-c & 1-c-d \\ 1 & 0 & 0 & 0 \\ 0 & 1-c-d & c & d \\ 0 & 1-b-c & b & c \end{pmatrix}$$

where a, b, c, d here denotes transition probabilities. The transitions U to U and W to W are assigned the same parameter since they represent the same configuration of pieces. Similarly, for the Markov chain along the diagonal some of the transitions are the same, and are identified with row transitions. In particular the diagonal transitions U to U and W to W are the same configuration as the row transition U to W (having probability b). In order to simplify the problem, we first assume that all diagonal elements in the transition matrix are the same, and we obtain

$$\mathbf{S} = \begin{pmatrix} b & e & c-e & 1-b-c \\ c & b & 1-b-c & 0 \\ 0 & 1-b-c & b & c \\ 1-b-c & c-e & e & b \end{pmatrix}.$$

Since the two matrices must commute (12), we compare the products of row 1 and column 4 and get the condition $d(c-e) = 0$. As the least restrictive choice we take $d = 0$. It is now possible to make the matrices commute by taking $b = (1-c)^2$, $e = c(2b+2c-1)$.

As an easy numerical case we take

$$\mathbf{R} = \begin{pmatrix} 0 & 1/4 & 1/4 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/4 & 1/4 & 1/2 \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/4 & 0 \\ 0 & 1/4 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

In this case all B, C combinations have probability $1/16$.

We can now find the entropy of the field by considering the conditional probability distribution of D given A, B, C . Clearly $C = L$ gives zero contribution to the entropy. Because of the choice $d = 0$, a W symbol cannot be followed by an U . Thus the

next symbol is either given as W because the one above was U , or it is not yet covered in which case L is the only possibility. Thus the entropy is also zero in this case. The two remaining cases, R and U , are identical. If $B = R$, D can be chosen as either L or U , and we get an entropy of 1 bit. For $B = U$, the only choices are U or W , depending on the symbol above, and the entropy is zero. For $B = L$ or W there are two situations: The following symbol is forced to be W with probability $1/4$ ($A = U$), or there are two choices for both A and D . In this case we can use iterative scaling (see Section III-B) or solve and optimize directly to get the marginals $1/4$ and $1/2$. Clearly, the conditional probability distribution is $(1/3, 2/3)$ in both cases. Thus, for these parameters we get the entropy

$$H = 1/8 + 3/16H(1/3) = 0.297.$$

With a more general version of \mathbf{S}

$$\mathbf{S} = \begin{pmatrix} b+c-e-f & e & f & 1-b-c \\ 2e+2f-c & b+c-e-f & 1-b-e-f & 0 \\ 0 & 1-b-c & b & c \\ 1-b-e-f & f & e & b \end{pmatrix}$$

we get 3 conditions on the variables by calculating the matrix products. These can be satisfied by selecting c, e , eliminating b, d as $b = 1 - 1.5c + (c^2 - f)/(2e + 2f)$ and $d = (c - e - f + ce + cf - c^2)/(c - e - 2f)$, and finally determining f as the solution of $-(1-c)^2 + b + d(1-c-e-b+f) = 0$. Now all admissible configurations have nonzero probabilities, and by varying the two parameters, we can get an entropy value greater than 0.35 for $c = 0.535$ and $e = 0.305$.

B. Iterative Scaling

Given two Markov chains having transition matrices which commute (12), the Pickard independence assumptions, $B \perp C | A$ and $B \perp C | D$, defines the two distributions (ABC) and (BCD) . The problem is now to define the probability distribution on the four elements $(ABCD)$, consistent with (ABC) and (BCD) and the constraint as expressed by (13)–(14).

Since the number of variables tends to be large, iterative scaling ([10, Th. 5.1]) may be used to find a solution, which also maximizes the entropy. Let a, b, c , and $d \in \mathcal{A}$ denote values of A, B, C , and D , respectively. Now let the values of $B = b$ and $C = c$ be fixed. Compute the marginal distribution on $cAb(P(A, B = b, C = c))$ and $cDb(P(D, B = b, C = c))$ from the Markov chains. The probability distribution $(AD|bc)$ can be written as a matrix, m_{ij} with the value of A defining the row index and the value of D defining the column index, i.e., for each configuration bc , the elements are given by

$$m_{ij} = P(A = i, D = j|bc), \quad (i, j) \in \mathcal{A}^2. \quad (15)$$

The row sums must equal the conditional distribution $(A|bc)$ (13) and the column sums $(D|bc)$ (14).

The iterative scaling may be described in terms of iterative projections. Following [10], the I -projection of a distribution $Q(x)$ onto a (nonempty) closed convex set Π is the $P^* \in \Pi$ such that $D(P^*||Q) = \min_{P \in \Pi} D(P||Q)$, where $D(P||Q) \equiv \sum P(x) \log(P(x)/Q(x))$.

The I-projection is easily found for linear families over partitions as defined, e.g., by either the row or the column sums in (13)–(14). Formally [10], the I-projection of $Q(x)$ onto a linear family \mathcal{L} determined by all distributions P having a given distribution $(\alpha_1, \dots, \alpha_k)$ on each of the k partitions (B_1, \dots, B_k) of A , is given by

$$P^*(a) = c_i Q(a), a \in B_i, \text{ where } c_i = \frac{\alpha_i}{Q(B_i)}. \quad (16)$$

Thus, P^* is obtained by scaling Q to \mathcal{L} .

In (13), the row and column sums each constitute a linear family over a partition. In general, consider m given linear families, $\mathcal{L}_1, \dots, \mathcal{L}_m$ and generate a sequence of distributions P_n by $P_0 = Q$ and for $n > 0$, let P_n be the I-projection of P_{n-1} onto \mathcal{L}_n , where for $n > m$, $\mathcal{L}_n = \mathcal{L}_i, i \equiv n \bmod m$. The sequence of distributions P_n converges to the I-projection onto the given linear families, ([10, Th. 5.1]),

Theorem 3.1: If $\cap_{i=1}^m \mathcal{L}_i = \mathcal{L} \neq \emptyset$ then $P_n \rightarrow P^*$, the I-projection of Q onto \mathcal{L} .

Let *iterative scaling* refer to the process of solving $P_n \rightarrow P^*$ by iterating scalings of the form (16).

Applying iterative scaling to (13)–(14), the matrix $\{m_{ij}\}$ is modified to get row sums $r_i = P(A = i|bc)$ (13) and column sums $c_j = P(D = j|bc)$ (14). Starting with the row sums defining \mathcal{L}_1 and mapping the initial $P_0 = Q$ and P_n for n even onto \mathcal{L}_1 by

$$m'_{ij} = m_{ij} * r_i / \sum_j (m_{ij}) \quad (17)$$

and subsequently, let the column sums define \mathcal{L}_2 and determine P_n for n odd by

$$m'_{ij} = m_{ij} * c_j / \sum_i (m_{ij}). \quad (18)$$

If the linear equations have a nonnegative solution, the iteration will converge to it by Theorem 3.1. For constraints defined on the joint variables, here A and D given bc , the matrix is initialized with zeros in positions that are not admissible by the constraint and this property is preserved by the scalings.

P^* is the solution, which minimizes the divergence, $D(P^*||Q)$ for $P^* \in \mathcal{L}$. Initializing with a uniform initial distribution over the admissible configurations as Q , implies that minimizing the divergence, $D(P||Q)$, by the iterative scaling (17)–(18) maximizes the entropy, $H(P)$, over the distributions $P \in \mathcal{L}$. We shall refer to this as *maximum entropy iterative scaling* and apply (17)–(18) to find a joint distribution on two variables given the marginal distribution on each variable.

It may be noted that, the initial matrix has the property that each entry (for an admissible configuration) satisfies

$$m_{ij} = t_i s_j \quad (19)$$

where t_i is a constant that depends on only the row index and s_j depends on only the column index. The iterative scaling

(17)–(18) preserves the form (19), as each step is a scaling of row or column values, as well as $m_{ij} = 0$ for non-admissible combinations.

C. Maximizing PRF Entropy for 2-D Constraints

Applying iterative scaling for each configuration of $B = b, C = c$ defines a distribution $(ABCD)$, which provides stationary solutions for the marginal distributions (ABC) and (BCD) satisfying the conditions of the Pickard random field (3), (4), (7), if such solutions exist. The *maximum entropy iterative scaling* of $P(AD|bc)$ is defined by (17) and (18) and selecting a uniform initial distribution over the admissible configurations. For each configuration bc , the elements are given by m_{ij} (15). The row and column sums (17)–(18) are given by $r_i = P(A = i|bc)$ and $c_j = P(D = j|bc)$ (13)–(14), respectively.

Iterative scaling is applied for each bc pair. The initial distribution, $Q(AD|BC)$, is set to a uniform distribution over the admissible configurations $abcd$ for each bc (and to 0 for the non-admissible configurations). For each bc a sequence of distributions is generated by iterating (17) and (18), which converges to the solution $P^*(AD|BC)$ in the limit. The entropy of the interior, $H(D|ABC)$, may be expressed by

$$\begin{aligned} H(D|ABC) &= H(ABCD) - H(ABC) \\ &= H(BC) + H(AD|BC) - H(ABC) \end{aligned} \quad (20)$$

where $H(BC)$ and $H(ABC)$ are given by the Markov chains defining the boundary and $H(AD|BC)$ is given by the distribution determined by the iterative scaling.

Theorem 3.2: It is a necessary condition for two irreducible Markov chains, with transition matrices \mathbf{R} and \mathbf{S} , to define the boundary (5) of a Pickard random field satisfying (7), that these matrices commute (12). For two irreducible Markov chains satisfying (12) and defining a boundary (5) by \mathbf{R} and \mathbf{S} , applying iterative scaling (17)–(18), to find a solution for $P(AD|BC)$ (13)–(14), determines whether a PRF with the given boundary and satisfying (4) and (7) exists. If a solution exists, the maximum entropy iterative scaling I-projections, $P^*(AD|bc)$ for all values of bc , combined with \mathbf{R}, \mathbf{S} (4) and (7), determines the conditional probabilities $P(D|ABC)$, which maximizes the entropy, $H(D|ABC)$, among PRFs with the boundary description (5) given by \mathbf{R} and \mathbf{S} . Initializing the iterative scalings with $Q(ad|bc) = 0$ for forbidden configurations $abcd$ of \mathcal{F} within a 2×2 square, the PRF will satisfy the constraint defined by \mathcal{F} .

Proof: The PRF independence conditions, $B \perp C | A$ (4) and $B \perp C | D$ (7), implies that the boundaries are defined by two Markov chains. For a given MC boundary pair (5), with stochastic matrices \mathbf{R} and \mathbf{S} , the PRF independence conditions (4) and (7) also define the distributions (ABC) (4) and (BCD) (7). The transition matrices for C to B over A and D are \mathbf{SR} and \mathbf{RS} , respectively. Starting with any given value c of C , the distribution of Bc must be identical for \mathbf{SR} and \mathbf{RS} . As the Markov chains are irreducible, this applies for all values of c and therefore $\mathbf{SR} = \mathbf{RS}$ is a necessary condition.

To ensure stationarity of the PRF, the stationary distributions of A, B, C and D must be identical. Thus the stationary distributions of \mathbf{R} and \mathbf{S} must be selected when defining (ABC) by (4) and (BCD) by (7). This in turn implies that the distributions (AB) and (CD) as well as (AC) and (BD) will be identical. Thus a distribution $(ABCD)$ consistent with the so defined (ABC) and (BCD) , will be stationary.

The stationary distribution (BC) as well as $P(A|bc)$ and $P(D|bc)$ (13)–(14) are given for all values of bc by \mathbf{R} and \mathbf{S} and the independence conditions (4) and (7). To complete the PRF, $P(D|ABC)$ must be specified. By Theorem 3.1, iterative scaling (17)–(18) based on (13)–(14) for all values of bc finds a solution to the joint distribution $P(AD|BC)$, consistent with $P(A|BC)$ (13) and $P(D|BC)$ (14), if such a solution exists.

Rewriting $H(D|ABC)$ by (20), the terms $H(BC)$ and $H(ABC)$ are fixed for given \mathbf{R} and \mathbf{S} . $H(AD|BC)$ expressed by $P(AD|bc)$ for each bc is the only variable term and it is optimized by the maximum entropy iterative scaling (17)–(18) based on (13)–(14), thus optimizing $H(D|ABC)$ among PRFs for which the boundary is given by \mathbf{R} and \mathbf{S} .

Initializing the iterative scalings with $Q(ad|bc) = 0$, for forbidden configurations, $abcd$, the iterative scalings (17)–(18) preserve this property and converges to the I-projection with $P^*(ad|bc) = 0$ for forbidden $abcd$, which in turn gives $P(abcd) = 0$ for forbidden $abcd$. Thus forbidden configurations appearing in the interior will be assigned probability 0. By the stationarity property of the PRF, this also applies to forbidden configurations appearing on the boundary. \square

Comment: Theorem 3.2 deals with PRF where the elements are traversed from the upper left hand corner (5) and condition $B \perp C \mid D$ (7) is satisfied. We can also consider the alternative pair of independence conditions, $A \perp D \mid B$ and $A \perp D \mid C$ (8), for defining a PRF in Theorem 2.2. In this case one of the Markov chains could be reversed and again the condition (12) could be checked. Again Theorem 3.2 could be applied but now switching the roles of B, C and A, D . One could also consider reversing both Markov chains, but this is already covered as reversing both irreducible MCs, the matrices commute iff \mathbf{R} and \mathbf{S} commute.

IV. BLOCK PICKARD RANDOM FIELDS FOR 2-D CONSTRAINTS

To construct fields with constraints of larger extent than 2×2 elements, we consider PRFs where the symbols are replaced by supersymbols defined by the symbols within rectangular blocks. The blocks are rectangles of size K by $2J$ symbols. We restrict the treatment to blocks arranged in rows in such a way that each row is shifted by half a block, J symbols, relative to the previous row (Fig. 2). The blocks are identified by two indices, the first one indicating the position on a diagonal from upper left to lower right. Thus block $g(i, j)$ includes rows $(i-1)K+1$ to iK . The other index increases from left to right.

We propose a causal model where the blocks are specified one diagonal at a time starting from the upper left and extending to the lower right. Each diagonal is given conditioned on the previous one. If we need to refer to a specific element, it will be given a third index numbered $1 \dots 2KJ$ row by row within each block.

$$\begin{array}{ccc} & & c_1 c_2 c_3 c_4 \\ & & c_5 c_6 c_7 c_8 \\ c_1 c_2 & & a_1 a_2 a_3 a_4 d_1 d_2 d_3 d_4 \\ a_1 a_2 d_1 d_2 & & a_5 a_6 a_7 a_8 d_5 d_6 d_7 d_8 \\ b_1 b_2 & & b_1 b_2 b_3 b_4 \\ & & b_5 b_6 b_7 b_8 \end{array}$$

Fig. 2. Four 1×2 blocks ($J = K = 1$) and four blocks with $J = K = 2$.

Now consider blocks A, B, C, D positioned on two successive diagonals with C above B (see Fig. 2) and consider N by M constraints. By choosing $K \geq N-1$, we can describe the sequences of blocks along the diagonals as Markov chains, since the forbidden configurations of \mathcal{F} involving elements of the current block does not extend beyond the rows of the previous block in the causal direction. By choosing $K \geq N-1$ and $J \geq M-1$, an $N \times M$ configuration that includes a symbol in D does not extend into past blocks other than A, B , and C . Further, for the chosen values of J and K , no forbidden N by M configuration involves elements of both B and C . Thus it is not impossible off-hand that they can be independent for all admissible configurations in a PRF. We base our construction of a block PRF on these blocks. As we shall see later for specific constraints, the properties of the forbidden configurations on the list \mathcal{F} allows for blocks sizes smaller than $K = N-1$ by $2J = 2M-2$ still allowing for independence of B and C for all admissible configurations.

Example 4.1: Consider a binary field with the constraint that allows at most one 1 within an N by $(M=N)$ square which may be expressed by

$$x_{ij} x_{i'j'} = 0, \quad \text{if } \max\{|i-i'|, |j-j'|\} < N.$$

Thus a list of forbidden configurations, \mathcal{F} , would include all combinations of two 1s within such a square. For $N = 3$ we can base the construction of a PRF on the 2×4 blocks indicated in Fig. 2. Here $N-1 = K = 2, M-1 = J = 2$ and the individual elements within a block are given by lower case letters. In the simpler case $N = 2$, we can use the same 1×2 blocks (Fig. 2) as for the n.i.b. field (Fig. 1). Let the indices of the blocks be given by the the binary number of the binary symbols within the block, e.g., $a = 2a_1 + a_2, a_1, a_2 \in \{0, 1\}$ (Fig. 2). We consult the list of forbidden configurations of \mathcal{F} both when defining the possible configurations of a block and the two blocks of a transition, this gives $a \in \{0, 1, 2\}$ as there can not be two 1s within the block. It is a straight forward exercise to show that we can get the transition matrix (A to B) for the Markov chain as

$$\mathbf{R} = \begin{pmatrix} \frac{1-3x}{1-x} & 1 & \frac{1-3x}{1-x} \\ \frac{x(1-3x)}{(1-x)(1-2x)} & 0 & \frac{2x}{1-x} \\ \frac{x}{1-2x} & 0 & 0 \end{pmatrix}$$

Here $P(01) = P(10) = x$ and the commuting transition matrix, \mathbf{S} , is obtained by interchanging the last two rows and columns. The conditional probabilities $P(D|ABC)$ may be obtained by iterative scaling by finding a solution to (17)–(18). In this example, the structure (19) of the maximum entropy

iterative scaling leads to a unique solution for a given value of x . The maximal entropy is $H = 0.4112$ per binary symbol, which is achieved for $x = 0.1304$. The entropy value is close to the capacity for the constraint, which was estimated to be 0.425 [11]. It may be noted that if the blocks were organized in an ordinary horizontal-vertical structure, then for the constraint elements of B and C would be diagonal neighbors leading to exclusion of admissible configurations to achieve the independence of B and C (given A) (4) regardless of the size of the blocks.

Now consider blocks A, B, C, D of $K \times 2J$ elements positioned together as given above with block D at $g(i, j)$ and consider this the current block. For a PRF defined on the blocks to satisfy a constraint it is obviously a necessary condition that configurations on A, B, C, D which are forbidden must be assigned probability 0.

Theorem 4.2: Given a list of forbidden configurations, \mathcal{F} , within an $N \times M$ rectangle and a 2-D lattice structure with rectangular blocks tiling the plane, consider a Pickard Random Field description on these blocks over the alphabet of the configurations on the blocks. The causal ordering of elements is given by (6) on the 2-D lattice and a given ordering of elements within the block. A sufficient condition for this block PRF to assign the probability 0 to any configuration containing a forbidden configuration of \mathcal{F} , is that the conditional probability is $P(D|ABC) = 0$, when the last symbol of the forbidden configuration of \mathcal{F} , in the causal ordering, is an element of the block D .

Proof: The condition states that any forbidden configuration of \mathcal{F} with the last causal element in D conditional on symbols in A, B, C (and D) will assign the configuration of $\mu_2 \times 2$ probability 0 (6). By the PRF stationarity this will also apply to all configurations (6) having at least one element of a forbidden configuration in the interior (6). Again applying the PRF stationarity that all rows as well as all columns have identical (marginal) distributions ensures that configurations with a forbidden configuration of \mathcal{F} is assigned probability 0 also on the boundary. \square

It should be noted that it is not always possible to construct a PRF in this way. There may be cases where there is no admissible choice for block D or it may not be possible to satisfy the independence constraints. The symmetric run-length constraints [13] is an example where issues of finding solutions is nontrivial. For $K \geq N - 1$ and $J \geq M - 1$, any N by M rectangle holding elements of D , will only have causal elements in the blocks A, B , and C and no forbidden configurations of \mathcal{F} has elements in both B and C . Below we consider specific constraints where the list of forbidden configurations \mathcal{F} allows smaller block sizes while still providing a restriction of checking the constraint to configurations of $ABCD$ and a separation of B and C .

A. No Isolated Bits

Block Pickard random field construction for 2-D constraints is now treated for the 2-D no isolated bits (n.i.b.) constraint (Example 1.1) in detail and with a slight modification the no isolated zero (n.i.z) constraint (where the isolated 1 is admissible

and only an isolated zero is forbidden). For these constraints $M = N = 3$, but the block Pickard field along diagonals can be constructed with 1×2 blocks as the forbidden configurations do not involve the corner elements, so the relation of the center element and the diagonal neighbors are not important. For the n.i.b. constraint, consider the four 1×2 blocks constituting $ABCD$ (Fig. 2). Given the values of A, B , and C , we must choose the values of D in such a way that neither a_2 or d_1 becomes isolated. It is not necessary to check other causal elements besides A, B , and C . Thus if we can define a PRF, it will assign probability 0 to any forbidden configuration by Theorem 3.2. It may also be observed that, the choice $d_1 = d_2 = a_2$ will always be admissible for any combination of elements of A, B , and C . This implies that each of the four configurations on the diagonal processes defines a state and all transitions along diagonals are possible as well as all configurations of ABC . This is also true for the less restrictive n.i.z. constraint.

A simple solution for the transition probability matrices to commute is simply to assign equiprobable transition probabilities to the transitions between the four symbols in the diagonal processes. A higher entropy may be achieved giving a bias to certain configurations.

Again the configurations on the blocks are indexed by the binary numbers, e.g., $a = 2a_1 + a_2, a_1, a_2 \in \{0, 1\}$. If we consider the transitions from C to A and C to D (see Fig. 2), it is clear that $(c_1 c_2) = (10)$ to (00) are two different configurations in the two processes. However, if we consider the transitions from A to B and C , the configurations are obtained by reflection in the horizontal axis, and we could choose to assign the same probability to them. Here AB is the same transition as CD , but AC is a transition in the same chain as CA taken in the reverse order. Thus in addition to symmetry around a vertical axis, and symmetry with respect to interchange of the binary symbols 0 and 1, we can simplify the construction by requiring that the two diagonal processes represent the same Markov chain with a reversal of the direction. (Unfortunately that property does not in itself imply that the transitions matrices commute.) If we write the transition matrix (with transitions from columns to rows) for the main diagonal (A to B) as

$$\mathbf{R} = \begin{pmatrix} aehd \\ bfgc \\ cgfb \\ dhea \end{pmatrix} \quad (21)$$

where the symmetry with respect to an interchange of 0 and 1 has been used to reduce the number of parameters. (a, b, c, d are here and in the next section used as parameters.) This property implies that the stationary distribution has the form $(1/2 - s, s, s, 1/2 - s)$. Since the columns of \mathbf{R} sum to 1, we have

$$b + c = 1 - (a + d) \quad (22)$$

$$f + g = 1 - (e + h) \quad (23)$$

and if s is used as a parameter, we get

$$(1/2 - s)(b + c) + s(f + g) = s \quad (24)$$

which gives

$$f + g = 1 - (b + c)(1 - 1/(2s)) \quad (25)$$

If the other transition matrix, \mathbf{S} , is specified by using symmetry about a vertical axis, we simply have to interchange rows and columns 2 and 3 of \mathbf{R} . To make this matrix the transition matrix for the process in the main diagonal reversed, we multiply the columns of both by the stationary probability of the corresponding states, and then require that one of the results is the transpose of the other. (This procedure amounts to verifying that $P(A = x, B = y) = P(A = y, B = x)$.) This condition actually implies just that

$$e/c = (1/2 - s)/s \quad (26)$$

since it follows from the previous relations that we then also have

$$h/b = e/c. \quad (27)$$

To make \mathbf{R} and \mathbf{S} commute, we calculate the product and notice that eight of the entries are already equal due to the symmetry of the matrices. The remaining eight products are equal if

$$(a - d - f + g)(b - c) = 0. \quad (28)$$

We prefer to make the first factor equal to zero, since this is less restrictive than $b = c$. This analysis leaves us with the parameters s, a, d , and b . For a set of parameters, iterative scaling (17)–(18) may be applied to obtain the conditional probabilities $P(D|ABC)$ thus giving the full parameter set of a block PRF. For each bc configuration the matrix with elements $m_{ij} = P(a(i)d(j)|bc)$ is determined using maximum entropy iterative scaling (17)–(18). Thereafter the entropy, $H(D|ABC)$, of the field can be calculated by (20). Starting with $s = 0.25, a = 0.2, d = 0.3, b = 0.2$, the entropy becomes 0.9091. By varying the initial choices of parameters, we find the maximum value $H(D|ABC) = 0.9157$ bits per binary symbol for $s = 0.2241, a = 0.2250, d = 0.3360$, and $b = 0.2457$. For comparison a lower bound of 0.9127 for bit-stuffing was given in [5] for the n.i.b. constraint. In a recent paper [12], a lower bound of 0.9226 was presented based on a larger neighborhood.

A block PRF may also be described for the no isolated zero (n.i.z.) constraint. All configurations admissible by the n.i.b. constraint are also admissible by the n.i.z. constraint. Using the same boundary model as for n.i.b. above, maximum entropy iterative scaling was again applied. Optimizing over the PRF boundary parameters gave $H(D|ABC) = 0.9559$ bits per binary symbol for the parameters $s = 0.236, a = 0.278, b = 0.243, d = 0.236$.

B. Minimum Distance 3 Between 1's

We construct a binary block Pickard field for the constraint that the 1-norm distance between 1's is at least 3 [9], which elements $x_{ij} \in \{0, 1\}$ may be defined as

$$x_{ij}x_{i'j'} = 0, \quad \text{if } |i - i'| + |j - j'| < 3. \quad (29)$$

Thus this constraint may be expressed by an intersection of constraints defined on pairs of elements. The basic constraint is described as a 3×3 square. However, since a pair of 1's positioned at diagonal corners of the 3×3 square is admissible (29), it is sufficient to use $J = 1, K = 2$, i.e., 2×2 blocks (shifted by one symbol in successive rows). This is sufficient to ensure that when selecting D given ABC , all causal elements with respect to D within distance 3 (29) of an element in D , and thereby all configurations of \mathcal{F} are in A, B or C . Therefore, if a PRF can be defined with $P(d|abc) = 0$ for $abcd$ forbidden, the block PRF will assign probability 0 to forbidden configurations (Theorem 4.2).

The states of the Markov chains on 2×2 blocks are defined by the configurations satisfying (29) within the block. This gives five states: the zero state and four states with a single 1. The nonzero states will be listed with the state that has a 1 in the upper left corner first, and the one with a 1 in the lower right corner last.

For each diagonal MC, the transitions are again defined by pairwise testing (29) of the elements of the two blocks. There are five transitions which are not admissible, leaving 15 variables (since the columns have to sum to 1).

To simplify the construction, we assume that all nonzero states have the same stationary probability, x . Because of the symmetry, the two transition matrices can be obtained by interchanging two pairs of rows and columns. The two transition matrices, \mathbf{R} from A to B and \mathbf{S} from C to A , are (with transitions from columns to rows)

$$\mathbf{R} = \begin{pmatrix} a & d & e & b & c \\ yc & g & 0 & 0 & 0 \\ yb & k & g & j & 0 \\ ye & h & i & g & 0 \\ yd & l & h & k & g \end{pmatrix} \quad (30)$$

$$\mathbf{S} = \begin{pmatrix} a & e & d & c & b \\ yb & g & k & 0 & j \\ yc & 0 & g & 0 & 0 \\ yd & h & l & g & k \\ ye & i & h & 0 & g \end{pmatrix} \quad (31)$$

where the parameters a, b, c, d, e follow from the sum of the columns, and the factor $y = x/(1 - 4x)$ in the first column serves to give the desired stationary distribution. It follows from the choice of parameters that probabilities of symmetric transitions are equal. However, for the matrices to commute, four pairs of products which correspond to transitions that are not symmetric, have to be equal. When \mathbf{S} is obtained by interchanging rows and columns of \mathbf{R} , the eigenvectors are found by applying the same permutation. It can be observed that g is an eigenvalue with eigenvector of the form $(0, 1, -1, -f, f)$. This observation gives $k = fj, i = h$, and $l = kf + i$. From direct calculation of the products we find

$$j = \frac{2y(1-g)(1+f)}{2f + y(1+f)^2} \quad (32)$$

and,

$$i = \frac{y(1-g)(1-f^2)}{1 + 2y(1+f)} \quad (33)$$

Thus for any value of $x \in [0, 1/4]$, the transition matrices are specified by the parameters g and f . When the transition matrices are given, we can fix the values of the blocks B and C , find the marginal distributions of the blocks A and D from the two Markov chains, and finally obtain the distribution $(ABCD)$ by the process of iterated scaling (17)–(18). We can then vary the parameters to get the maximal entropy of the field. By searching over the parameters, the maximum entropy is found to be $H(D|ABC) = 0.3341$ for the parameters $x = 0.0987$, $f = k/j = 0.5102$, $g = 0.1567$. This value is somewhat lower than the best lower bound for the constraint, 0.3503 [9]. This lower bound was based on a model, which is not finite, conditioning on the causal part of the last N rows.

V. CONCLUSION

Block Pickard random fields were introduced. With a suitable two-variable indexing of the blocks along diagonals a probability assignment is derived from Markov chain descriptions. If the transition matrices of the two Markov chains can be chosen to commute and the iterative scaling process has a solution for the conditional distribution of a new block, the field is described as a block Pickard random field. In this case details of the distribution can be explicitly calculated, and in particular we calculate the entropy. Block PRF were calculated for a number of 2-D constraints.

REFERENCES

- [1] S. Friedland, "On the entropy of Z^d subshifts of finite type," *Linear Algebra Its Appl.*, vol. 252, pp. 199–220, 1997.
- [2] J. Justesen and Yu. M. Shtarkov, "The combinatorial entropy of images," (in Russian) *Probl. Pered. Inf.*, vol. 33, pp. 3–11, 1997.
- [3] A. Kato and K. Zeger, "On the capacity of two-dimensional run-length constrained channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1527–1540, Jul. 1999.
- [4] R. M. Roth, P. H. Siegel, and J. K. Wolf, "Efficient coding scheme for the hard-square model," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1166–1176, Mar. 2001.
- [5] S. Halevy, J. Chen, R. M. Roth, P. H. Siegel, and J. K. Wolf, "Improved bit-stuffing bounds on two-dimensional constraints," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 824–838, May 2004.
- [6] D. Pickard, "Unilateral Markov fields," *Adv. Appl. Probab.*, vol. 12, pp. 655–671, 1980.
- [7] F. Champagnat, J. Idier, and Y. Goussard, "Stationary Markov random fields on a finite rectangular lattice," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2901–2916, Nov. 1998.
- [8] R. Burton and R. Pemantle, "Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances," *Ann. Probab.*, vol. 21, pp. 1329–1371, 1993.
- [9] S. Forchhammer and J. Justesen, "Entropy bounds for constrained two-dimensional random fields," *IEEE Trans. Inf. Theory*, vol. 45, pp. 118–127, Jan. 1999.
- [10] I. Csiszár and P. C. Shields, "Information theory and statistics: A tutorial," *Found. Trends in Commun. Inf. Theory*, vol. 1, no. 4, pp. 422–527, 2004.
- [11] W. Weeks and R. E. Blahut, "The capacity and coding gain of certain checkerboard codes," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1193–1203, Jan. 1999.
- [12] I. Tal and R. M. Roth, "Bounds on the rate of 2-D bit-stuffing encoders," in *Proc. ISIT*, Toronto, Canada, 2008, pp. 1463–1467.
- [13] T. Etzion, "Cascading methods for run-length limited arrays," *IEEE Trans. Inf. Theory*, vol. 43, pp. 319–324, Jan. 1997.

Søren Forchhammer (M'04) received the M.S. degree in engineering and the Ph.D. degree from the Technical University of Denmark, Lyngby, in 1984 and 1988, respectively.

Currently, he is an Associate Professor with DTU Fotonik, Technical University of Denmark, where he has been since 1988, also as Head of the Coding Group. His main interests include source coding, image and video coding, distributed source coding, two-dimensional information theory, two-dimensional fields, and visual communications.

Jørn Justesen (S'68–M'98) was born in Copenhagen, Denmark, in 1944. He received the M.S. degree in electrical engineering in 1971 and the Dr.Tech. degree in 1975, both from the Technical University of Denmark, Lyngby.

From 1971 to 1972, he was with the Institute of Circuit Theory and Telecommunication, Technical University of Denmark. He was a Visiting Lecturer with the University of California, Berkeley, in 1973, and an Associate Professor with Aalborg University, Denmark, from 1974 to 1975. Since 1976, he has been Professor of Communication Theory in the Department of Telecommunication (now Photonics), Technical University of Denmark. His main research interests are algebraic coding theory and constrained codes in one and two dimensions. In fall 1987, he visited the Institute for Problems of Information Transmission, Moscow, and visited the University of Maryland, College Park, in spring 1999. He has worked on several contracts for the European Space Agency.

Dr. Justesen received the Browder J. Thompson Prize for the best paper published in any IEEE TRANSACTIONS in 1972 by an author under 30 years old. He was the coauthor of the paper that received the IEEE Information Theory 1991 Best Paper Award. He was Guest Associate Editor for a special issue of the IEEE TRANSACTIONS ON INFORMATION THEORY in 1995, and an Associate Editor for coding theory of the TRANSACTIONS in 2000–2003. In 2004, he was awarded the title of Professor Honoris Causa by the Institute for Problems of Information Transmission, Moscow.